Field Crops Research 216 (2018) 32–41

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Whole-genome prediction of reaction norms to environmental stress in bread wheat (*Triticum aestivum* L.) by genomic random regression

Delphine Ly[a], Sylvie Huet[b], Arnaud Gauffreteau[c], Renaud Rincent[a], Gaëtan Touzy[d], Agathe Mini[d], Jean-Luc Jannink[e,f], Fabien Cormier[g], Etienne Paux[a], Stéphane Lafarge[d], Jacques Le Gouis[a], Gilles Charmet[a,*]

[a] *Institut National de la Recherche Agronomique-Clermont-Auvergne University, UMR1095, Genetics, Diversity and Ecophysiology of Cereals, 63000 Clermont-Ferrand, France*
[b] *Institut National de la Recherche Agronomique, MAIAGE – Applied Mathematics for Computer Sciences, Jouy-en-Josas, France*
[c] *Institut National de la Recherche Agronomique, UMR Agronomy, Thiverval Grignon, France*
[d] *Biogemma, Research Center, Route d'Ennezat, Chappes, France*
[e] *USDA-ARS, R.W. Holley Center, 538 Tower Rd, United States*
[f] *Cornell University, Section of Plant Breeding and Genetics, Ithaca, NY, United States*
[g] *CIRAD, UPR Multiplication Végétative, Avenue Agropolis, TA-A108/03, 34398 Montpellier Cedex 5, France*

ARTICLE INFO

Keywords:
Genotype-by-environment interaction
Factorial regression
Genomic prediction
Reaction norm
Drought adaptation

ABSTRACT

Plant breeding has always sought to develop crops able to withstand environmental stresses, but this is all the more urgent now as climate change is affecting the agricultural regions of the world. It is currently difficult to screen genetic material to determine how well a crop will tolerate various stresses. Multi-environment trials (MET) which include a particular stress condition could be used to train a genomic selection model thanks to molecular marker information that is now readily available. Our study focuses on understanding how and predicting whether a plant is adapted to a particular environmental stress. We propose a way to use genomic random regression, an extension of factorial regression, to model the reaction norms of a genotype to an environmental stress: the factorial regression genomic best linear unbiased predictor (FR-gBLUP). Twenty-eight wheat trials in France (3 years, 12 locations, nitrogen or water stress treatments) were split into two METs where different stresses limited grain number and yield. In MET1, drought at flowering was responsible for 46.7% of the genotype-by-environment (G × E) interactions for yield while in MET2, heat stress during booting was identified as the main factor responsible for G × E interactions, but that explained less of the interaction variance (33.6%).

Since drought at flowering explained a fairly large variance in G × E in MET1, the FR-gBLUP model was more accurate than the additive gBLUP across all types of cross validation. Accuracy gains varied from 2.4% to 12.9% for the genomic regression to drought. In MET2 accuracy gains were modest, varying from −5.7% to 2.4%. When a major stress influencing G × E is identified, the FR-gBLUP strategy makes it possible to predict the level of adaptation of genotyped individuals to varying stress intensities, and thus to select them *in silico*. Our study demonstrates how genome-wide selection can facilitate breeding for adaptation.

1. Introduction

A single plant genotype can produce a multiplicity of phenotypes in response to the environment. This phenotypic plasticity reflects the plant's ability to sense, respond to and survive a diversity of abiotic stresses (Bohnert et al., 1995; Des Marais et al., 2013). One way to assess the importance of genetic variation in plasticity is to measure the genotype-by-environment interaction (G × E). When selecting genotypes for breeding, it is complicated to deal with large G × E interactions particularly when the ranking of the genotypes changes across environments (Haldane, 1946). The definition of a superior individual is thus conditional on the environment in which the individuals are

Abbreviations: gBLUP, best linear unbiased predictor; FR-gBLUP, factorial regression best linear unbiased predictor; G × E, genotype-by-environment; Sum of P + I-ETP < 0, sum of negative hydric balance modeled by the difference between rainfall and potential evapotranspiration; SNP, single nucleotide polymorphism; CVrandom, random cross-validation; CVnewE, cross-validation in new environments; CVnewG, cross-validation for new genotypes; CVnewGE, cross-validation for new genotypes in new environments

* Corresponding author.
E-mail address: gilles.charmet@inra.fr (G. Charmet).

https://doi.org/10.1016/j.fcr.2017.08.020
Received 15 December 2016; Received in revised form 19 August 2017; Accepted 25 August 2017
Available online 11 November 2017
0378-4290/

tested, and thus, conditional on the stresses the genotype might encounter. Climate change is already causing wheat yields to stagnate in Europe (Brisson et al., 2010), so we need greater insight into the genetics of plant adaptation to the environment (Via et al., 1995).

Reaction norms have been widely used to statistically model how genotypes are adapted to some measure of the environment quality, calculated by joint regression (Finlay and Wilkinson, 1963; Fischer and Maurer, 1978; Lin and Binns, 1988) or factorial regression (Denis, 1980; Denis et al., 1997; van Eeuwijk et al., 1996). In factorial regression models the G × E interaction term represents the sensitivities of each genotype to environmental stress covariates. An advantage of factorial regression is the potential to predict untested environments or local adaptation (Smith et al., 2005). Ravagnolo and Misztal (2000) proposed a random regression model to model genotypic sensitivities of dairy cattle to heat stress using a pedigree-based relationship. With the recent development of molecular markers, factorial regression has been applied to predict untested genotypes by modeling different effects of alleles at a QTL when regressed on environmental covariates (Malosetti et al., 2004). In wheat trials, a multi-environment QTL model has identified QTL by environment interaction for drought stress at anthesis, underlining the fact that marker alleles may have different effects in different environments (Mathews et al., 2008). With faster, cheaper and more comprehensive genome-wide marker information available, genomic selection has been developed to use all markers across the genome simultaneously to predict the breeding values of quantitative traits (Meuwissen et al., 2001).

Genomic selection models have mostly been developed to predict additive genetic values, thus ignoring interactions with environments. Recently, some genomic models have been proposed to extend the genomic best linear unbiased predictor (BLUP) to take G × E interaction into account (Burgueño et al., 2012; Crossa et al., 2013; Heslot et al., 2014; Jarquín et al., 2014; Saint Pierre et al., 2016). Because the interaction is likely to be caused by several environmental stresses, these models focused on handling the multiplicity of stress covariates. However, models for genomic predictions of G × E have not yet exploited the predictive ability of factorial regression to model the differential responses to a particular environmental stress.

In this paper we explore the environmental stresses that influenced the G × E interactions in a trial network of 28 environments and develop a genomic model to predict the reaction norms to a specific environmental factor. Such predictions can then be used to select varieties having a specific response to those environmental factors, and thus select adaptations to environments where these factors are likely of influence the trait of interest. We illustrate the usefulness of the model in bread wheat by analyzing the response of grain number to water deficit stress. Finally, we discuss the use of this model in breeding for adaptation in wheat.

2. Materials and methods

2.1. Experimental data

Two hundred twenty-one European wheat varieties were evaluated in 3 years (2012, 2013 and 2014) at 12 locations in France (Cappelle, Châlons-en-Champagne, Gréoux-les-Bains, Réalville, Clermont-Ferrand, Estrées-Mons, Verneuil l'Etang, Champigny, Louville, Maule, Andelu, Vraux) (Fig. 1). Drought experiments were conducted in each site-year combination at Gréoux-les-Bains, Réalville and Champigny with either water deficit (DRY) or irrigated conditions (IRR). For the other site-year combinations two nitrogen (N) treatments were applied, either low (N−) or high (N+). In total, 28 combinations of year, location and stress treatment (N deficiency or drought) were studied in this experiment. Grain yield and yield components were measured for the 221 varieties. Different block designs were used at each site but all had the same four check genotypes common to each block (i.e. augmented design). Data from each trial was corrected for block effects estimated based on the common check genotypes, using a linear model with genotype as a random effect. An autoregressive process in the direction of each row and column was used to model the spatial trend in trials if necessary. For this, variograms were checked and the Aikaike information criterion (AIC), Bayesian information criterion (BIC) and likelihood ratio test were used to select the best model.

Climatic data for each trial was also available. Inter-annual and inter-latitude variations covered a wide range of environmental conditions in France. Environments were described using monthly climatic variables (radiation, sum of average temperature, sum of rainfall, sum of negative rainfall minus potential evapotranspiration, sum of temperatures above 25 °C) coinciding approximately with the developmental stages of wheat (April, booting; May, meiosis and flowering; June, grain filling). In particular, water stress was measured as the sum of rainfall and irrigation minus the potential evapotranspiration, to reflect the balance between plant water demand and supply (Allen et al., 1998; Doorenbos and Pruitt, 1977). Drought at anthesis is known to impact spike fertility, grain number and yield (Farooq et al., 2009; Fischer and Maurer, 1978; Gupta et al., 2001) and heat stress during the booting stage in April and May, e.g. measured as the sum of temperatures above 25 °C, has been shown to strongly impact grain number, and thereby yield (Farooq et al., 2011; Gate, 1995; Porter and Gawith, 1999).

All the lines were genotyped with an Affymetrix Axiom 280 K SNP array developed in the framework of the BreedWheat project (Rimbert et al., in preparation). The markers producing missing values, values of heterozygosity above 0.05, or showing monomorphism were discarded from the analysis. Redundant markers were also discarded by using a linkage disequilibrium threshold of $r^2 = 0.9$, computed chromosome by chromosome. This resulted in the selection of 77369 markers for further analyses.

2.2. BLUP predictions of a genomic random regression to an environmental stress, mixed-model factorial regression genomic BLUP (FR-gBLUP)

We describe here a genomic random regression to an environmental stress covariate model, that we called factorial regression genomic BLUP model or FR-gBLUP. We modeled here the genomic random regression (b_G) to the environmental covariate x_E, so that the reaction norms to x_E were more similar for related than for unrelated genotypes.

$$y = \mu 1_n + \beta Z_E x_E + Z_G u_G + (Z_G b_G) \circ (Z_E x_E) + Z_E u_E + e \quad \text{(Model 1)}$$

$$\text{With} \begin{pmatrix} \mathbf{u_G} \\ \mathbf{b_G} \\ \mathbf{u_E} \\ \mathbf{e} \end{pmatrix} \sim N \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_G^2 \mathbf{K} & \sigma_{GB} \mathbf{K} & \mathbf{0} & \mathbf{0} \\ \sigma_{GB} \mathbf{K} & \sigma_B^2 \mathbf{K} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_E^2 \mathbf{I}_{nE} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \sigma_{nE}^2 \end{pmatrix} \otimes \mathbf{I_{nG}} \end{pmatrix} \right)$$

In this model y is the adjusted mean for grain number at the plot level corrected by the block effect; μ is the overall mean; x_E is the vector of the environmental covariate with a single covariate value per environment, centered on 0 and scaled to a variance of 1; β is the coefficient of the fixed regression to x_E; u_G represents the genetic effects following a normal distribution $u_G \sim N(0, \sigma_G^2 K)$ where K is the genomic relationship matrix calculated using VanRaden (2008) formula, b_G is the coefficient of the random regression to the environmental covariate x_E for each genotype and represents the specific genotypic linear reaction norm to the environmental covariate x_E following the normal distribution $b_G \sim N(0, \sigma_B^2 K)$; the symbol ∘ denotes the Hadamard product; σ_{GB}^2 is the covariance between the random effects of the intercepts (u_G) and the slopes (b_G); u_E is the random effect of the environments; Z_G and Z_E are the incidence matrices for the genotypic and environment effects respectively; I represents the identity matrix; nE is the number of environments; n_G is the number of observations; and e represents the heteroscedastic residuals with an environment specific error variance.

The phenotypic trait y is thus modeled as a linear function of the environmental covariate x_E, with u_G being the intercept and b_G a

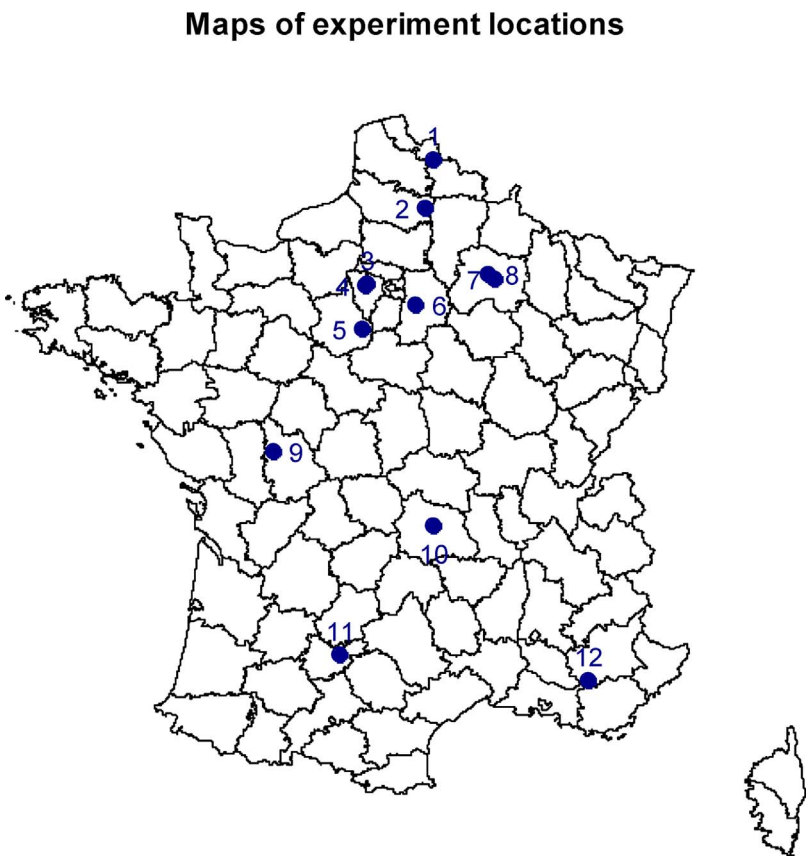

Fig. 1. Map ofexperiment locations.: (1) Cappelle, (2) Estrées-Mons, (3) Andelu, (4) Maule, (5) Louville, (6) Verneuil, (7) Vraux, (8) Châlons-en-Champagne, (9) Champigny, (10) Clermont-Ferrand, (11) Réalville, (12) Gréoux-les-Bains.

random deviation to the average slope β depending on the genotypes. The different slopes b_G of the genotypes are modeled considering their genomic relationships. An example of Model 1 is described in Appendix A.

We compared this model to the additive genomic BLUP model (Hayes et al., 2009; Habier et al., 2013), which does not include the $b_G * x_E$ terms. For more details on the additive model and its prediction equations, see Appendix B. For more details on the prediction equations of the FR-gBLUP model, see Appendix C.

The variance components were estimated using Restricted Maximum Likelihood. We computed these models using ASReml 3 in R (Gilmour et al., 2006).

2.3. Assessment of the prediction quality of the models

The prediction quality of the different models was assessed by cross validations. We used 4 different cross validation schemes: CVrandom, CVnewG, CVnewE, and CVnewGE. All cross validations consisted of a 4-fold cross validation (Kohavi, 1995), using 75% of the data in the training set and 25% in the validation set. The variance components were estimated for each training set. In CVrandom, the folds were built randomly from the observations of all individuals in different environments. CVnewG consisted of building the folds so that there were no common genotypes between the training set and the validation set. Thus, CVnewG simulated situations where a model is built using some genotypes to predict the performance of new (untested) lines in the same set of environments. Similarly, CVnewE had no common environments, i.e. location × year climatic conditions between training and validation sets. Thus CVnewE simulated situations where the performances of tested genotypes can be predicted in new environments. Finally, CVnewGE split the data according to the genotypes and the environment, so that predictions were made for new genotypes in new environmental conditions. The cross-validations were repeated 10 times.

To evaluate the usefulness of modeling an interaction term, we calculated the Pearson correlation cor(y,$\hat{y}$) within location × year × treatment environments, where $\hat{y} = \hat{u}_G$ without interaction (Model in Appendix B), and $\hat{y} = \hat{u}_G + \hat{b}_G x_E$ with interaction (Model 1).

3. Results

3.1. Agro-climatic indicators reveal a diverse panel of trial environments

The experiment sites covered a wide range of French latitudes (Fig. 1), and climatic conditions were also diverse during the years of the experiment. It should be noted that the north of France was more represented than the south, which coincides with the amount of land area given over to wheat crops in the different regions. In terms of radiation, water stress and temperature, environments generally showed similar behavior during April. Differences become apparent in May and June as variations in these climatic variables grew wider (Fig. 2). Gréoux environments particularly became clearly differentiated from all other environments. Irrigation and nitrogen treatments also contributed to the broad diversity of environments. Within the same site-year conditions, water stress treatments produced yields which were less well correlated than between N treatments (Fig. 3).

We clustered the environments into two METs. MET1 consisted of dry and irrigated environments at Gréoux in 2012, 2013 and 2014, while MET2 included all the other trials. Indeed, climatic data showed that the environmental conditions in MET1 (in Gréoux) contrasted with MET2 (Fig. 2). A hierarchical clustering of genotypic yield values between environments clearly distinguished the two METs. Yields were negatively correlated between MET1 and MET2 (Fig. 3). Clustering the environments into two groups according to grain number and thousand kernel weight, showed that the Gréoux environments were still associated to each other (results not shown). The limiting stress factors and

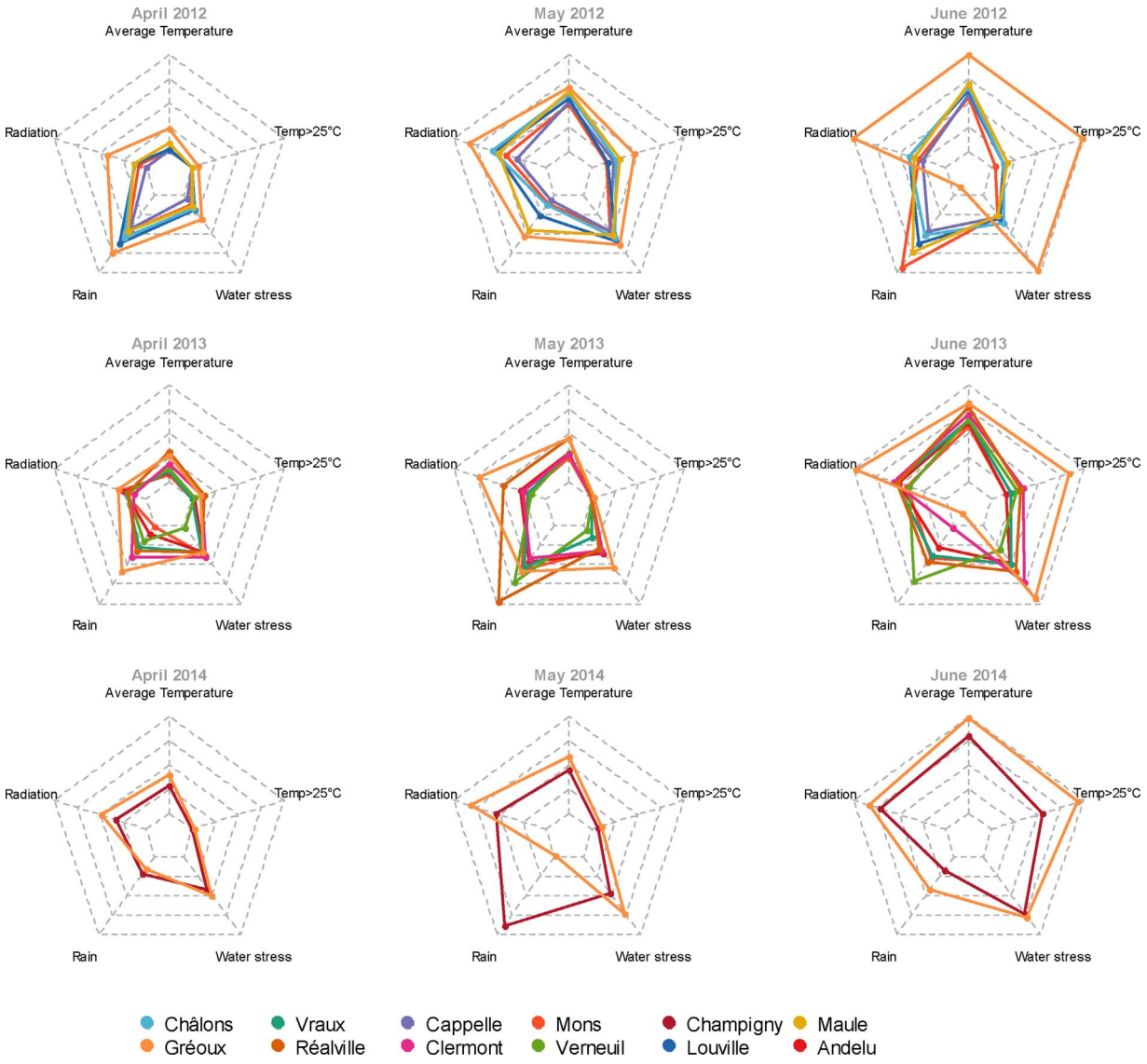

Fig. 2. Characterization of environments according to 5 environmental variables (average temperature, temperatures above 25 °C, radiation, rain, water stress) from April to June in 2012, 2013 and 2014. The series of radar charts represent different growing seasons (rows) and growth years (columns). Within each chart, the colored lines (one for each environment) show how similar they are.

also probably the causes of GxE interactions were likely to be different for each of the trial network.

For the purpose of evaluating the predictive ability of the FR-gBLUP model, we identified drought at flowering as the covariate of interest in MET1, and heat in June in MET2. Indeed, drought at flowering in MET1 had the highest variance of reaction norms (Table 2 for MET1). In MET2, the variance of reaction norms to heat in April May only ranked 2nd (Supp. Data 1 for MET2). However contrary to the stress with the highest variance, (i) it affected only the traits that were thought to be impacted by the timing of the stress (grain number and yield), (ii) the direction of the slopes showed that the performances of the genotypes were decreased with the stress intensity, and more importantly (iii) the distribution of the environmental covariate was uniform and not structured by years.

3.2. Water stress around flowering causes high G × E interactions in Gréoux environments

MET1 displayed higher G × E interaction variability (Table 1). The G × E interaction variance represented 53% to 75% of the genotypic variance in yield for MET1 compared to 0% to 64% in MET2, and 19% to 44% of the genotypic variance in grain number for MET1, but 0% to 42% for MET2. In MET1, water stress in May was identified as the major stress causing G × E interaction affecting yield, as it showed the highest variance in the reaction norms (Table 2). Besides, it also generated high variance of reaction norms for grain number. In the other environments, despite lower reaction norms variance than in MET1, the major stress causing G × E interaction was heat stress during booting (Supplementary Data 1).

3.3. Modeling the genomic random regression to water stress improves prediction accuracies

The FR-gBLUP was used to model the genomic reaction norms to water stress in May in MET1. The results showed an improvement in the prediction of grain number and yield compared to gBLUP. The average gain in prediction accuracy varied from 2.4% to 12.9% for yield and from 3.3% to 8.0% for grain number depending on the type of cross-validation applied (Table 3 & Fig. 4). In MET2, where no single major stress influenced G × E interactions as strongly as in MET1, modeling the reaction norms to heat stress during booting resulted in rather modest gains for CVrandom, 1.0% for grain number and 2.4% for yield. There was even some loss of accuracy for all other cross-validations from −0.6% to −2.7% for grain number and −0.6% to −5.7% for yield (Table 3).

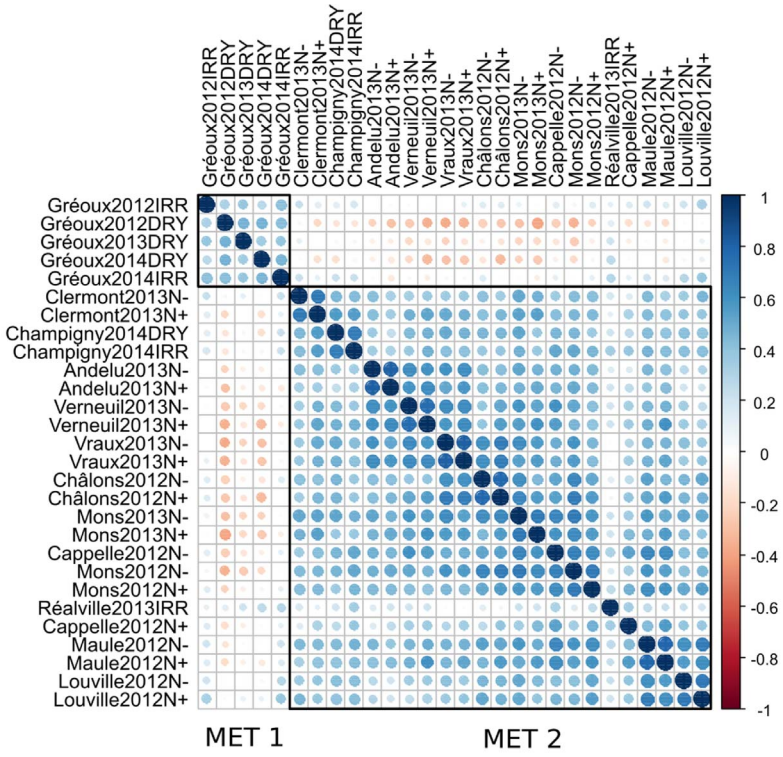

Fig. 3. Correlation of the genotypic values for wheat yield between trials. A trial is defined as a location × year × treatment environment. A hierarchical clustering based on these correlations defined multi-environment trials MET 1 and MET 2. See Materials and Methods for details of trial environments.

4. Discussion

4.1. Identifying the major environmental stress factors causing G × E interactions

Here we present two case studies attempting to identify the major environmental stress(es) involved in interactions between a wheat genotype and its environment. Correlation of the genotypic values for wheat yield between environments clearly distinguished two trial networks, MET1 and MET2 (Fig. 3), indicating that the limiting stress factors were likely to be different between the groups. The decision to split the dataset was motivated by the negative correlations of genotypic values for yield observed between groups (Fig. 3) and to avoid inflating prediction accuracies in Gréoux due to an environmental covariate which would mainly distinguish Gréoux environments from all others. As MET1 was composed of the Gréoux drought trials from

Table 1
Variance of the reaction norms to different environmental covariates in Gréoux environments. These variances were estimated separately in reaction norm models where genotypes were assumed to be independent. The variance of GxE represent the total interaction variance that we aim to capture through an environmental covariate. P refers to rainfall, I to irrigation, PET to potential evapotranspiration, T to temperature. Yield mean per environment was tested.

Effect		Grain Number	Thousand Kernel Weight	Yield
Genotype		2.74 ± 0.31	0.12 ± 0.01	0.24 ± 0.03
Environment		13.33 ± 9.44	0.21 ± 0.15	3.56 ± 2.52
GxE		1.06 ± 0.12	0.02 ± 0.003	0.09 ± 0.02
	Environmental Covariate			
	Sum Radiation May	0.42 +/− 0.253	0.019 +/− 0.01	0 +/− 0
	Sum Tmax > 25 June	0.658 +/− 0.211	0.009 +/− 0.007	0 +/− 0
	Deficit Radiation June	1.466 +/− 0.155	0 +/− 0	0 +/− 0
	Sum Radiation June	0.682 +/− 0.127	0.005 +/− 0.002	0.001 +/− 0.011
	Sum Radiation Meiosis Flowering	0.104 +/− 0.204	0.019 +/− 0.008	0.001 +/− 0.019
	Sum Tmax > 25 April–May	0.182 +/− 0.03	0.001 +/− 0	0.002 +/− 0.002
	Sum Tmax > 25 May	0.208 +/− 0.034	0.001 +/− 0	0.002 +/− 0.002
	Average T June	0.138 +/− 0.071	0.006 +/− 0.003	0.004 +/− 0.006
	Average T May	0.663 +/− 0.139	0.009 +/− 0.003	0.007 +/− 0.009
	Deficit Radiation May	0.005 +/− 0.048	0.006 +/− 0.002	0.009 +/− 0.006
	Sum Rain April	0.203 +/− 0.037	0.001 +/− 0	0.01 +/− 0.004
	Sum Rain June	0.313 +/− 0.062	0.003 +/− 0.001	0.011 +/− 0.006
	Sum Radiation April	0 +/− 0	0.008 +/− 0.003	0.013 +/− 0.008
	Excess Water June	0.332 +/− 0.064	0.003 +/− 0.001	0.014 +/− 0.007
	Sum P + I-ETP June	0.112 +/− 0.04	0.003 +/− 0.001	0.016 +/− 0.006
	Deficit Radiation April	0.558 +/− 0.139	0.012 +/− 0.004	0.017 +/− 0.012
	Sum P + I-ETP April	0.506 +/− 0.078	0.002 +/− 0.001	0.019 +/− 0.008
	Sum Rain May	0.195 +/− 0.048	0.003 +/− 0.001	0.024 +/− 0.008
	Average T April	0.22 +/− 0.075	0.005 +/− 0.001	0.028 +/− 0.011
	Sum Tmax > 25 April	0.394 +/− 0.138	0.008 +/− 0.003	0.037 +/− 0.017
	Sum P + I-ETP May	0.441 +/− 0.083	0.003 +/− 0.001	0.042 +/− 0.012
Residuals		2.32 ± 0.10	0.05 ± 0.002	0.47 ± 0.02

Table 2
Decomposition of grain number and yield variance in a mixed linear model y = G + E + G × E + e where genotypes are assumed to be independent and where G × E is modeled with a different G × E variance per environment.

		Grain Number	Yield
Genotype		3.64 ± 0.35	0.32 ± 0.03
Environment		13.98 ± 3.81	2.04 ± 0.56
Genotype × Environment			
	Châlons2012N-	0 ± 0	0.01 ± 0.02
	Châlons2012N +	0.04 ± 0.08	0.03 ± 0.02
	Gréoux2012IRR	1.03 ± 1.87	0.36 ± 0.05
	Gréoux2012DRY	2.07 ± 3.76	0.93 ± 0.11
	Gréoux2013DRY	1.65 ± 3	0.82 ± 0.09
	Gréoux2014IRR	0.84 ± 1.52	0.63 ± 0.07
	Gréoux2014DRY	2.9 ± 5.26	0.96 ± 0.11
	Vraux2013N-	0.39 ± 0.71	0.02 ± 0.01
	Vraux2013N +	0.61 ± 1.1	0.03 ± 0.01
	Réalville2013IRR	1.29 ± 2.35	0.42 ± 0.05
	Cappelle2012N-	0.27 ± 0.49	0.09 ± 0.02
	Cappelle2012N +	0.83 ± 1.51	0.24 ± 0.03
	Clermont2013N-	0.33 ± 0.59	0.16 ± 0.03
	Clermont2013N +	0.53 ± 0.97	0.11 ± 0.02
	Mons2012N-	0.01 ± 0.01	0.05 ± 0.02
	Mons2012N +	0.24 ± 0.44	0.07 ± 0.02
	Mons2013N-	0 ± 0	0 ± 0
	Mons2013N +	0.17 ± 0.31	0.1 ± 0.02
	Verneuil2013N-	0.63 ± 1.14	0.11 ± 0.02
	Verneuil2013N +	2.64 ± 4.8	0.18 ± 0.03
	Champigny2014IRR	0.25 ± 0.45	0.16 ± 0.03
	Champigny2014DRY	0.17 ± 0.31	0.15 ± 0.02
	Louville2012N-	1.6 ± 2.9	0.52 ± 0.06
	Louville2012N +	2.02 ± 3.68	0.64 ± 0.07
	Maule2012N-	0.88 ± 1.61	0.23 ± 0.03
	Maule2012N +	0.99 ± 1.81	0.35 ± 0.04
	Andelu2013N-	1.15 ± 2.09	0.28 ± 0.04
	Andelu2013N +	1.86 ± 3.39	0.36 ± 0.05
Residuals		1.82 ± 0.03	0.21 ± 0

2012 to 2014, it was expected from the experimental design that drought would indeed be identified as a major stress. This is a straightforward case of identifying the stress generating G × E interactions. Indeed, the regression to water deficit stress around flowering accounted for 41.6% of the G × E interaction for grain number and 46.7% for yield (Table 2). It is still very useful to check whether the stress imposed in the experiment is not just generating a large E effect, but also whether there really are G × E effects, i.e. stress responses that vary from one genotype to another. If this is not the case, then the stress will not generate much G × E, even though it may be affecting the trait in a similar manner for all genotypes. In MET2, which assembled all the environments except the ones in Gréoux, most trials showed low or no water deficit stress. In MET2, heat stress during booting was the main stress causing G × E but this response represented a smaller proportion of the G × E variance, 25.6% for grain number and 33.6% for yield (Supplementary Data 1). That is why we chose to illustrate the use of FR-gBLUP on the MET1 data.

To identify an environmental stress as potentially responsible for G × E interactions, the response to this stress should first explain a large enough proportion of the G × E interaction but should also make sense. i) The direction of reaction norms should agree with what is known of the eco-physiology, and ii) the traits thought to be impacted by the timing of the stress (e.g., grain number and yield for stress during booting) should have a large variance of response and meaningful reaction norms. In this way more confidence can be put into the identification of an environmental covariate generating G × E interactions. For the analysis of MET2, these are reasons why we rejected the sum of temperatures above 25 °C in June and chose the sum of temperatures above 25 °C in April and May.

For further work, defining environmental stress per actual stage of development would provide more accurate environmental covariates. Especially with such a gradient of latitude in MET2, the stress characterization by monthly statistics may not fit well the real phenological stages. For example, crop models could help focusing on a particular stage to compute stress indices (Heslot et al., 2014).

4.2. Prediction of the genomic reaction norms to an environmental covariate can be made only for a major stress influencing G × E interactions

By incorporating environmental stress covariates into the classical genomic BLUP model, our objective was to be able to predict the reaction norms to a stress for unobserved genotypes. We studied 2 cases that suggest that in making genomic predictions of reaction norms, the environmental covariate has to be a major stress influencing G × E interactions. Prediction accuracy gains were achieved for water stress as the environmental covariate, when the reaction norms explained 40% of the total G × E interactions.

The response to water stress around flowering showed a fairly large variance in MET1, as high as 17.2% of the genetic additive variance for grain number and 16.7% for yield, indicating that there is an interaction signal which could be modeled to improve genomic predictions (Table 2). Indeed the predictions made with the FR-gBLUP model gained in accuracy (Table 3) compared to the additive model. On the other hand, for temperature stress, when the proportion of the variance was only 9% of the additive variance (although 33.6% of the G × E variance), no gain in accuracy was achieved with any of the cross validation strategies. For genomic models, one of the most difficult challenges is to predict the performance of new (untested) lines, a major objective for breeders. The ability to do this depends on the strength of the genetic relationship between genotypes in the training population and genotypes in the testing or target population (Habier et al., 2010, 2007; Ly et al., 2013; Rincent et al., 2012). In our study, we found a consistent decrease in accuracy, as well as increasing variation in prediction accuracies between replicates, from CVrandom to CVnewG and CVnewGE (Fig. 4). This confirms the difficulty of predicting new genotypes, depending on the relationship between reference individuals and predicted individuals. However, when modeling the response to water deficit stress around flowering, the FR-gBLUP model displayed noticeable gains in prediction accuracies for CVnewG (8.0% for grain number and 3.4% for yield) and CVnewGE (3.6% for grain number and 12.9% for yield). Thus, when predicting the performance of untested individuals, information on the reaction norms to an environmental stress covariate from relatives actually contributes to improving prediction accuracy through the FR-gBLUP model.

Table 3
Prediction accuracy gains of the FR-gBLUP relative to the additive gBLUP.

Environment	Environmental covariate	Cross-Validation	Accuracy Gain	
			Grain Number	Yield
MET1	Sum P + I-ETP May	CVrandom	3.3%	2.4%
		CVnewG	8.0%	3.4%
		CVnewE	3.3%	3.8%
		CVnewGE	3.6%	12.9%
MET2	SumTmax25ApMay	CVrandom	1.0%	2.4%
		CVnewG	−2.7%	−5.7%
		CVnewE	−0.6%	−0.6%
		CVnewGE	−0.6%	−3.2%

Another aspect of the FR-gBLUP model is that it addresses the challenge of predicting the behavior of genotyped individuals in new environments. Under CVnewE and CVnewGE, which require predictions for environments that have not been evaluated in the training set, FR-gBLUP exhibited moderate to relatively high gains in accuracy compared to the additive gBLUP model (Table 3). Consequently, while

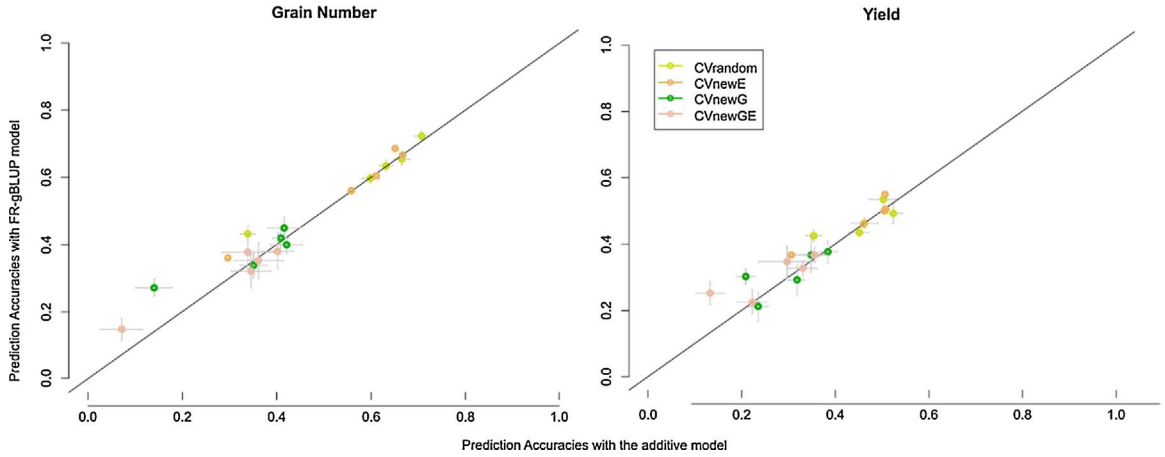

Fig. 4. Comparison of prediction accuracies between the FR-gBLUP and the additive models for responses to water stress in Gréoux environments for four cross-validation (CV) methods: random, or simulating prediction of unknown genotypes (newG), new environments (newE) or both (newGE). Points above the line indicate that the proposed model is more accurate than the baseline model.

modeling a genomic kinship for the reaction norms to an environmental stress covariate allows prediction of untested genotypes, the factorial regression model equips FR-gBLUP with some capacity to predict for untested environments.

Of course, the model we proposed for predicting the genotypic reaction norm to an environmental stress has some limitations. Factorial regression hypothesizes that the response to environmental covariates is linear in the parameters, which may not always mirror the actual response. In terms of selecting environmental variables, it is possible to add more than one environmental covariate as long as they are not too closely correlated to each other and that the degrees of freedom allowed by the dataset are not exceeded.

4.3. A step toward exploiting genomic G × E interactions in genomic selection

In our study, we did not aim at explaining the overall G × E interaction using both markers and environmental covariates. Rather we aimed at dissecting, and possibly predicting, the genetic response to a given stress. Therefore FR-gBLUP cannot be claimed to explain G × E as well as other models designed to predict overall G × E, and which usually include many environmental covariates, such as those proposed by Burgueño et al. (2012), Heslot et al. (2014) and Jarquín et al. (2014).

It is nevertheless instructive to compare FR-gBLUP to such models. For example, Jarquín et al. (2014) obtained a 34% gain in prediction accuracies with a cross-validation similar to our CVrandom, and a 17% gain with a cross-validation similar to CVnewG. Heslot et al. (2014) also modeled genomic G × E interactions to predict yield, for which G × E variance accounted for 63% of the total genetic variance. With a cross-validation similar to CVnewGE, they measured a gain of 11.1% in prediction accuracy, a lesser gain than in our study. We did not achieve gains as high as the those described by Jarquín et al. (2014), but with CVnewGE we obtained a prediction accuracy gain of 3.6% for grain number and 12.9% for yield by modeling the genomic regression to water deficit stress at flowering. The FR-gBLUP model thus showed its ability to improve prediction accuracy of untested genotypes in untested environments. These three previously published studies may not be directly comparable to ours because some uncertainty remains regarding (i) the amount of cross-over in the G × E interactions, (ii) the strength of the additive relationships in the population, (iii) the genotyping method, and finally (iv) the prediction accuracy of the additive model. Despite all these caveats, the FR-gBLUP model seems to be as good as some other G × E models at predicting G × E interactions when an appropriate covariate can be identified.

4.4. Applications for breeders and consultants

As an extension of the genomic BLUP, the FR-gBLUP makes it possible to predict the performance of new candidates based on their genotypic data (CVnewG). This is the main objective of breeders when they use genomic prediction instead of evaluation data to save time and money in their breeding schemes. It would be even more desirable to predict the behavior of new genotypes in new environments, but this approach would cumulate two sources of uncertainties.

Nowadays the need to accelerate breeding for yield potential and better adaptation to the changing climate is a critical issue. Adaptation is a complex trait to predict, likely determined by many genes and their interactions. The main approach developed is to gain more insight into the physiology of the tolerance to environmental stress to define physiological traits and use genetics to explain and predict these traits. The FR-gBLUP takes advantage of the huge amount of marker data available to predict the response to the environmental stress as a selection trait of interest. Genomic selection has the potential to accelerate selection cycles, and may go beyond breeding for adaptation to stress to breeding for better use of environmental resources when crops are not stressed.

5. Conclusions

We have developed a genomic random regression extension of the factorial regression model, called FR-gBLUP, to predict the reaction norms of a crop to an environmental stress covariate. In this case study of wheat, we identified an environmental stress which explains a large proportion of the G × E interactions. For this condition, the FR-gBLUP model performed better than an additive model. The FR-gBLUP model also provides insights into the understanding of G × E and broadens the choice of genotypes that can be recommended to withstand particular environmental stresses or used for adaptation breeding.

Authors' contributions

DL, SH, AG and GC designed the analysis. GT and AM collected the data. GT, AM, RR, DL and GC analyzed the phenotypic data. DL and SH developed the models and the equations. DL, SH, AG, FC, RR, JLJ and GC interpreted and discussed the results. DL performed the statistical analyses and wrote the manuscript with the help of AG, GC, RR.

Funding

The experimental work was supported by the BreedWheat project thanks to funding from the French Government managed by the

National Research Agency (ANR) in the framework of Investments for the Future (ANR-10-BTBR-03), France AgriMer and the French Fund to support Plant Breeding (FSOV). This work was supported by the INRA metaprogram SELGEN and the Auvergne Region, by funding the doctorate of DL.

Competing interest

The authors declare they have no competing interests.

Acknowledgements

We are grateful to Sue Welham and the ASReml support for their help. We also thank David Gouache for the constructive and enriching discussions. The authors thank the reviewers for their comments, which helped us improve the manuscript.

Appendix A. Example of the FR-gBLUP (Model 1)

$$y = \mu 1_n + \beta Z_E x_E + Z_G u_G + (Z_G b_G) \circ (Z_E x_E) + Z_E u_E + e \text{ (Model 1)}$$

$$\text{with} \begin{pmatrix} \mathbf{u}_G \\ b_G \\ u_E \\ e \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_G^2 K & \sigma_{GB} K & \mathbf{0} & \mathbf{0} \\ \sigma_{GB} K & \sigma_B^2 K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_E^2 I_{nE} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & \sigma_{nE}^2 \end{pmatrix} \otimes I_{\mathbf{nE}} \end{pmatrix} \right)$$

Let us consider 3 genotypes G1, G2 and G3 evaluated in 2 environments E1 and E2. Observations of their phenotypes are written as $y_{Gi,Ej}$. The environments E1 and E2 are characterized by an environmental stress covariate x_{Ej}.

Let $y = \begin{pmatrix} y_{G1,E1} \\ y_{G2,E1} \\ y_{G3,E1} \\ y_{G1,E1} \\ y_{G2,E2} \\ y_{G3,E2} \end{pmatrix}$ be the vector of the phenotypic observations. In this case, $Z_E x_E$ is $\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{E1} \\ x_{E2} \end{pmatrix}$, $Z_G u_G$ is equal to $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_{G1} \\ u_{G2} \\ u_{G3} \end{pmatrix}$, and similarly, the genotypic response term can be written as $Z_G b_G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_{G1} \\ b_{G2} \\ b_{G3} \end{pmatrix}$

With these notations, Model 1 becomes:

$$\begin{pmatrix} y_{G1,E1} \\ y_{G2,E1} \\ y_{G3,E1} \\ y_{G1,E2} \\ y_{G2,E2} \\ y_{G3,E2} \end{pmatrix} = \mu \mathbf{1}_n + \beta \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{E1} \\ x_{E2} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_{G1} \\ u_{G2} \\ u_{G3} \end{pmatrix} + \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_{G1} \\ b_{G2} \\ b_{G3} \end{pmatrix} \right) \circ \left(\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{E1} \\ x_{E2} \end{pmatrix} \right) + Z_E u_E + e$$

$$\Leftrightarrow \begin{pmatrix} y_{G1,E1} \\ y_{G2,E1} \\ y_{G3,E1} \\ y_{G1,E2} \\ y_{G2,E2} \\ y_{G3,E2} \end{pmatrix} = \mu \mathbf{1}_n + \beta \begin{pmatrix} x_{E1} \\ x_{E1} \\ x_{E1} \\ x_{E2} \\ x_{E2} \\ x_{E2} \end{pmatrix} + \begin{pmatrix} u_{G1} \\ u_{G2} \\ u_{G3} \\ u_{G1} \\ u_{G2} \\ u_{G3} \end{pmatrix} + \begin{pmatrix} b_{G1} \\ b_{G2} \\ b_{G3} \\ b_{G1} \\ b_{G2} \\ b_{G3} \end{pmatrix} \circ \begin{pmatrix} x_{E1} \\ x_{E1} \\ x_{E1} \\ x_{E2} \\ x_{E2} \\ x_{E2} \end{pmatrix} + Z_E u_E + e$$

$$\Leftrightarrow \begin{pmatrix} y_{G1,E1} \\ y_{G2,E1} \\ y_{G3,E1} \\ y_{G1,E2} \\ y_{G2,E2} \\ y_{G3,E2} \end{pmatrix} = \mu \mathbf{1}_n + \beta \begin{pmatrix} x_{E1} \\ x_{E1} \\ x_{E1} \\ x_{E2} \\ x_{E2} \\ x_{E2} \end{pmatrix} + \begin{pmatrix} u_{G1} \\ u_{G2} \\ u_{G3} \\ u_{G1} \\ u_{G2} \\ u_{G3} \end{pmatrix} + \begin{pmatrix} b_{G1} x_{E1} \\ b_{G2} x_{E1} \\ b_{G3} x_{E1} \\ b_{G1} x_{E2} \\ b_{G2} x_{E2} \\ b_{G3} x_{E2} \end{pmatrix} + Z_E u_E + e$$

Appendix B. Additive genomic BLUP model

The additive gBLUP model consisted of modeling the genetic additive effects of the genotypes as a random effect with a variance covariance structured by a genomic estimate of kinship between the individuals.

$$y = \mu \mathbf{1}_n + \beta Z_E x_E + Z_G u_G + Z_E u_E + e$$

$$\text{With}\begin{pmatrix}\mathbf{u}_G\\ u_E\\ e\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}\sigma_G^2 K & 0 & 0\\ 0 & \sigma_E^2 I_{nE} & 0\\ 0 & 0 & \begin{pmatrix}\sigma_1^2 & 0 & 0\\ 0 & \ldots & 0\\ 0 & 0 & \sigma_{nE}^2\end{pmatrix}\otimes I_{nE}\end{pmatrix}\right)$$

In the additive model, y is the adjusted mean for grain number at the plot level corrected by the block effect; μ is the overall mean; x_E is the vector of the environmental covariate, with a single covariate value per environment, centered on 0 and scaled to a variance of 1; β is the coefficient of the fixed regression to x_E; u_G is the genetic effects following a normal distribution $u_G \sim N(0, \sigma_G^2 K)$ with K being the genomic relationship matrix calculated using VanRaden (2008) formula; u_E is the random effect of the environments; u_{GE} is the random effect of the genotype-by-environment interactions; Z_G, Z_E, Z_{GE} are the incidence matrices for the genotype, environment or genotype-by-environment effects respectively; I and represents the identity matrix, nE the number of environments, n the number of observations and e the residuals.

According to the equation of prediction of random effects from Henderson (1975), the equation of the predictions of the genetic additive effects can be written

$$\hat{\mathbf{u}}_G = \hat{\sigma}_G^2 K Z'_G \hat{\mathbf{V}}^{-1}(y - \hat{\mu}\mathbf{1}_n - \hat{\beta} Z_E x_E) \tag{A.1}$$

$$\text{With}\mathbf{V} = \sigma_G^2 Z_G K Z'_G + \sigma_E^2 Z_E Z'_E + \begin{pmatrix}\sigma_1^2 & 0 & 0\\ 0 & \ldots & 0\\ 0 & 0 & \sigma_{nE}^2\end{pmatrix}\otimes I$$

To predict the performance $\hat{u}_{G\_new}$ of a new genotype, its marker information M_{G_new} was used to calculate its relatedness to the other genotypes. Thus the prediction of the additive effect for a new genotype was

$$\hat{\mathbf{u}}_{G.\mathbf{new}} = \hat{\sigma}_G^2\left(M_{G.new} M'/_k\right) Z'_G \hat{\mathbf{V}}^{-1}(y - \hat{\mu}\mathbf{1}_n - \hat{\beta} Z_E x_E) \tag{A.2}$$

Appendix C. Prediction equations of the genomic reaction norms

We calculated the BLUP predictions of the genomic additive effects and regression to x_E as:

$$\begin{pmatrix}\hat{\mathbf{u}}_G\\ \hat{\mathbf{b}}_G\end{pmatrix} = \begin{pmatrix}\hat{\sigma}_G^2 K & \hat{\sigma}_{GB} K\\ \hat{\sigma}_{GB} K & \hat{\sigma}_B^2 K\end{pmatrix}\begin{pmatrix}Z'_G\\ Z'_G\end{pmatrix}\hat{\mathbf{V}}^{-1}(y - \hat{\mu}\mathbf{1}_n - \hat{\beta} Z_E x_E) \tag{B.1}$$

$$\mathbf{V} = \sigma_G^2 Z_G K Z'_G + \sigma_{GB}(D Z_G K Z'_G + Z_G K Z'_G D') + \sigma_B^2 D Z_G K Z'_G D'$$

$$\text{With} \quad + \sigma_E^2 Z_E Z'_E + \begin{pmatrix}\sigma_1^2 & 0 & 0\\ 0 & \ldots & 0\\ 0 & 0 & \sigma_{nE}^2\end{pmatrix}\otimes I$$

where D is a diagonal matrix whose dimension is the number of observations, and which contains on the diagonal the elements of $Z_E u_E$.

Similarly, we wrote the predictions for genomic additive and interactive effects in model (2) for a new genotype as:

$$\begin{pmatrix}\hat{\mathbf{u}}_{G.\mathbf{new}}\\ \hat{\mathbf{b}}_{G.\mathbf{new}}\end{pmatrix} = \begin{pmatrix}\hat{\sigma}_G^2 M_{G.\mathbf{new}} M'/_k & \hat{\sigma}_{GB} M_{G.\mathbf{new}} M'/_k\\ \hat{\sigma}_{GB} M_{G.\mathbf{new}} M'/_k & \hat{\sigma}_B^2 M_{G.\mathbf{new}} M'/_k\end{pmatrix}\begin{pmatrix}Z'_G\\ Z'_G\end{pmatrix}\hat{\mathbf{V}}^{-1}(y - \hat{\mu}\mathbf{1}_n - \hat{\beta} Z_E x_E) \tag{B.2}$$

Appendix D. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fcr.2017.08.020.

References

Allen, R., Pereira, L., Raes, D., Smith, M., 1998. Crop evapotranspiration: guidelines for computing crop water requirements. In: FAO Irrigation and Drainage. Rome.

Bohnert, H., Nelson, D., Jensen, R., 1995. Adaptations to environmental stresses. Plant Cell 7, 1099–1111. http://dx.doi.org/10.1105/tpc.7.7.1099.

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X., Huard, F., 2010. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. F. Crop. Res. 119, 201–212. http://dx.doi.org/10.1016/j.fcr.2010.07.012.

Burgueño, J., de los Campos, G., Weigel, K., Crossa, J., 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci. 52, 707. http://dx.doi.org/10.2135/cropsci2011.06.0299.

Crossa, J., Vargas, M., Cossani, C.M., Alvarado, G., Burgueño, J., Mathews, K.L., Reynolds, M.P., 2013. Evaluation and interpretation of interactions. Agron. J. http://dx.doi.org/10.2134/agronj2012.0491.

Denis, J.B., Piepho, H.P., van Eeuwijk, F., 1997. Modelling expectation and variance for genotype by environment data. Heredity (Edinb) 79, 162–171.

Denis, J.B., 1980. Analyse de la régression factorielle. Biom. Praxim. 20, 1–34.

Des Marais, D.L., Hernandez, K.M., Juenger, T.E., 2013. Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. Annu. Rev. Ecol. Evol. Syst. 44, 5–29. http://dx.doi.org/10.1146/annurev-ecolsys-110512-135806.

Doorenbos, J., Pruitt, W.O., 1977. Crop water requirements. In: FAO Irrigation and Drainage. Rome.

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., Basra, S.M.A., 2009. Plant drought stress: effects, mechanisms and management. In: Lichtfouse, E., Navarrete, M., Debaeke, P., Véronique, S., Alberola, C. (Eds.), Sustainable Agriculture. Springer Netherlands, Dordrecht, pp. 153–188. http://dx.doi.org/10.1007/978-90-481-2666-8_12.

Farooq, M., Bramley, H., Palta, J.A., Siddique, K.H.M., 2011. Heat stress in wheat during reproductive and grain-filling phases. CRC Crit. Rev. Plant Sci. 30, 491–507. http://dx.doi.org/10.1080/07352689.2011.615687.

Finlay, K.W., Wilkinson, G.N., 1963. The analysis of adaptation in a plant-breeding program. Aust. J. Agric. Res. 14, 742–754. http://dx.doi.org/10.1071/AR9630742.

Fischer, R.A., Maurer, R., 1978. Drought resistance in spring wheat cultivars. I. Grain yield responses. Aust. J. Agric. Res. 29, 897–912. http://dx.doi.org/10.1071/AR9780897.

Gate, P., 1995. Ecophysiologie du blé. Technique et Documentation Lavoisier.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., 2006. ASReml User Guide Release

2.0. Hemel Hempstead, HP 1 1ES, UK.
Gupta, N., Gupta, S., Kumar, A., 2001. Effect of water stress on physiological attributes and their relationship with growth and yield of wheat cultivars at different stages. J. Agron. Crop Sci. 186, 55–62.
Habier, D., Fernando, R., Dekkers, J., 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177, 2389–2397. http://dx.doi.org/10.1534/genetics.107.081190.
Habier, D., Tetens, J., Seefried, F., Lichtner, P., Thaller, G., 2010. The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genet. Sel. Evol. 42, 5. http://dx.doi.org/10.1186/1297-9686-42-5.
Habier, D., Fernando, R.L., Garrick, D.J., 2013. Genomic BLUP decoded: a look into the black box of genomic prediction. Genetics 194, 597–607. http://dx.doi.org/10.1534/genetics.113.152207.
Haldane, J.B.S., 1946. The interaction of nature ad nurture. Ann. Hum. Genet. 13, 197–205. http://dx.doi.org/10.1111/j.1469-1809.1946.tb02358.x.
Hayes, B.J., Bowman, P.J., Chamberlain, a J., Goddard, M.E., 2009. Invited review: genomic selection in dairy cattle: progress and challenges. J. Dairy Sci. 92, 433–443. http://dx.doi.org/10.3168/jds.2008-1646.
Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31, 423. http://dx.doi.org/10.2307/2529430.
Heslot, N., Akdemir, D., Sorrells, M.E., Jannink, J.-L., 2014. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 127, 463–480. http://dx.doi.org/10.1007/s00122-013-2231-5.
Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., de los Campos, G., 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 127, 595–607. http://dx.doi.org/10.1007/s00122-013-2243-1.
Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence.
Lin, C.S., Binns, M.R., 1988. A superiority measure of cultivar performance for cultivar x location data. Can. J. Plant Sci. 68, 193–198.
Ly, D., Hamblin, M., Rabbi, I., Melaku, G., Bakare, M., Gauch, H.G., Okechukwu, R., Dixon, A.G.O., Kulakow, P., Jannink, J.-L., 2013. Relatedness and genotype × environment interaction affect prediction accuracies in genomic selection: a study in cassava. Crop Sci. 53, 1312. http://dx.doi.org/10.2135/cropsci2012.11.0653.
Malosetti, M., Voltas, J., Romagosa, I., Ullrich, S.E., Van Eeuwijk, F.A., 2004. Mixed models including environmental covariables for studying QTL by environment interaction. Euphytica 137, 139–145.
Mathews, K.L., Malosetti, M., Chapman, S., McIntyre, L., Reynolds, M., Shorter, R., van Eeuwijk, F., 2008. Multi-environment QTL mixed models for drought stress adaptation in wheat. Theor. Appl. Genet. 117, 1077–1091. http://dx.doi.org/10.1007/s00122-008-0846-8.
Meuwissen, T.H., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.
Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. Eur. J. Agron. 10, 23–36. http://dx.doi.org/10.1016/S1161-0301(98)00047-1.
Ravagnolo, O., Misztal, I., 2000. Genetic component of heat stress in dairy cattle, parameter estimation. J. Dairy Sci. 83, 2126–2130. http://dx.doi.org/10.3168/jds.S0022-0302(00)75095-8.
Rincent, R., Laloë, D., Nicolas, S., Altmann, T., Brunel, D., Revilla, P., Rodríguez, V.M., Moreno-Gonzalez, J., Melchinger, a, Bauer, E., Schoen, C.-C., Meyer, N., Giauffret, C., Bauland, C., Jamin, P., Laborde, J., Monod, H., Flament, P., Charcosset, a, Moreau, L., 2012. Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of maize inbreds (Zea mays L.). Genetics 192, 715–728. http://dx.doi.org/10.1534/genetics.112.141473.
Saint Pierre, C., Burgueño, J., Crossa, J., Fuentes Dávila, G., Figueroa López, P., Solís Moya, E., Ireta Moreno, J., Hernández Muela, V.M., Zamora Villa, V.M., Vikram, P., Mathews, K., Sansaloni, C., Sehgal, D., Jarquin, D., Wenzl, P., Singh, S., 2016. Genomic prediction models for grain yield of spring bread wheat in diverse agro-ecological zones. Sci. Rep. 6, 27312. http://dx.doi.org/10.1038/srep27312.
Smith, A.B., Cullis, B.R., Thompson, R., 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. J. Agric. Sci. 143, 449. http://dx.doi.org/10.1017/S0021859605005587.
VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91, 4414–4423. http://dx.doi.org/10.3168/jds.2007-0980.
Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S.M., Schlichting, C.D., Van Tienderen, P.H., 1995. Adaptive phenotypic plasticity: consensus and controversy. Trends Ecol. Evol. 10, 212–217.
van Eeuwijk, F.A., Denis, J.B., Kang, M.S., 1996. Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: Kang, M.S., Gauch, H.G. (Eds.), Genotype-by-Environment Interaction, pp. 15–50 (Boca Raton).